

Holonomic gradient method for the probability content of a simplex region with a multivariate normal distribution

Tamio Koyama

Abstract

We use the holonomic gradient method to evaluate the probability content of a simplex region with a multivariate normal distribution. This probability equals to the integral of the probability density function of the multivariate Gaussian distribution on the simplex region. For this purpose, we generalize the inclusion–exclusion identity which was given for polyhedra, to the faces of a polyhedron. This extended inclusion–exclusion identity enables us to calculate the derivatives of the function associated with the probability content of a polyhedron in general position. we show that these derivatives can be written as integrals of the faces of the polyhedron.

1 Introduction

The holonomic gradient method (HGM) is an algorithm for the numerical calculation of holonomic functions. It is a variation on the holonomic gradient descent (HGD) proposed in [8]. A holonomic function is an analytic function of several variables which satisfies a holonomic system. Here, a holonomic system refers to a system of linear differential equations with polynomial coefficients which induces a holonomic module in terms of D -module theory [9]. The HGM evaluates a holonomic function by numerically solving an initial value problem for an ordinary differential equation. This ordinary differential equation is derived from the Pfaffian equation (an integrable connection) associated with the function. For details, see [3] and its references. Several normalizing constants and the probability content of a region can be regarded as a holonomic function with respect to their parameters, and we can use the HGM to estimate the solution to this function. For example, the HGM was used to evaluate the cumulative distribution function for the largest root of the Wishart matrix in [2], and we utilized the HGM for the orthant probability and the Fisher–Bingham integral in [4] and [5], respectively.

Our motivation is to apply the HGM to the numerical calculation of the probability content of a polyhedron with a multivariate normal distribution. A polyhedron is a subset of a d -dimensional Euclidean space \mathbf{R}^d which is defined by

a finite number of linear inequalities. Intervals of real numbers, orthants, cubes, and simplexes are examples of polyhedra. In [7], Naiman and Wynn described examples of how the evaluation of the probability content of a polyhedron can be used to find critical probabilities for multiple comparisons.

The probability content of a polyhedron is a generalization of orthant probabilities discussed in [4] since the orthant probabilities can be expressed as the probability content of a simplicial cone. We derived the HGM for orthant probabilities in [4], and its implementation in [6]. A study of phylogenomics utilized our implementation in [11].

In order to utilize the HGM for the numerical calculation of the probability content of a polyhedron, we need to provide the Pfaffian equation explicitly and to evaluate the initial value for the ordinary differential equation. Our previous paper [3] showed that the probability content of a polyhedron in general position can be expressed as an analytic function and we explicitly provided a holonomic system and Pfaffian equations for this function. In this paper, we calculate the derivatives of the function, and show that these derivatives can be written as integrals of the faces of the polyhedron. This result provides formulae to compute the initial value exactly for the cases where the polyhedron is in general position and bounded, or the polyhedron is a simplicial cone.

In order to calculate the derivatives, we generalize the inclusion–exclusion identity that was given for polyhedra in [1], to the faces of the polyhedron. Since a face of a polyhedron is also a polyhedron, the inclusion–exclusion identity for the face holds, i.e., the indicator function of the face can be written as a linear combination of indicator functions of simplicial cones. Our generalized inclusion–exclusion identity gives the expression for this linear combination explicitly.

In the numerical experiments, we consider simplexes and simplicial cones, as these are fundamental examples of polyhedra. Utilizing the theoretical results concerned with the Pfaffian equation and the initial value, we implement the HGM to evaluate the probability contents of a simplex and a simplicial cone. We show that our implementation works well for a 10-dimensional simplex.

This paper is organized as follows. In section 2 we review results from our previous paper [3]. In section 3 we extend the inclusion–exclusion identity which was given for polyhedra in [1], and provide an analogous formula for the indicator function of a face of a polyhedron. In section 4 we calculate the derivatives of the function defined by the probability content of a polyhedron for the multivariate normal distribution, and show that these derivatives can be written by integrals on corresponding faces. In section 5 attention is directed towards the case where the polyhedron in general position is bounded, and the case where the polyhedron is a simplicial cone. We discuss the evaluation of the multivariate normal probabilities of polyhedra by the HGM in these two cases. In section 6 we present numerical examples.

Acknowledgements. This work was supported by JSPS KAKENHI Grant Number 263125.

2 Summary of Previous Work

In this section we review the results of our previous paper [3]. Let us consider a polyhedron

$$P := \left\{ x \in \mathbf{R}^d : \sum_{i=1}^d \tilde{a}_{ij} x_i + \tilde{b}_j \geq 0, 1 \leq j \leq n \right\} \quad (1)$$

where $\tilde{a}_{ij}, \tilde{b}_j$ ($1 \leq i \leq d, 1 \leq j \leq n$) are real numbers. We denote by \tilde{a} and \tilde{b} the $d \times n$ matrix (\tilde{a}_{ij}) and the vector $(\tilde{b}_1, \dots, \tilde{b}_n)^\top$ respectively. We suppose that the polyhedron P is in general position and its bounding half-spaces are

$$H_j := \left\{ x \in \mathbf{R}^d : \sum_{i=1}^d \tilde{a}_{ij} x_i + \tilde{b}_j \geq 0 \right\} \quad (1 \leq j \leq n). \quad (2)$$

For the definitions of general position and the set of the bounding half-spaces for a polyhedron, see [3].

We denote by \mathcal{F} the abstract simplicial complex associated with the polyhedron P , i.e.,

$$\mathcal{F} := \left\{ J \subset \{1, 2, \dots, n\} \mid \bigcap_{j \in J} H_j \neq \emptyset \right\}.$$

Let

$$\varphi(a, b) = \int_{\mathbf{R}^d} \frac{1}{(2\pi)^{d/2}} \exp \left(-\frac{1}{2} \sum_{i=1}^d x_i^2 \right) \sum_{F \in \mathcal{F}} \prod_{j \in F} (H(f_j(a, b, x)) - 1) dx \quad (3)$$

be a function with variables a_{ij}, b_j ($1 \leq i \leq d, 1 \leq j \leq n$). Here, $H(x)$ is the Heaviside function and we set $f_j(a, b, x) = \sum_{i=1}^d a_{ij} x_i + b_j$. We denote by a and a_j ($j = 1, \dots, n$) the $d \times n$ matrix with elements a_{ij} , ($i = 1, \dots, d$) and the column vector $(a_{1j}, \dots, a_{dj})^\top$ respectively. And b is a column vector $(b_1, \dots, b_n)^\top$ with length n . For $J \in \mathcal{F}$, we put

$$g^J(a, b) = \left(\prod_{j \in J} \partial_{b_j} \right) \bullet \varphi(a, b). \quad (4)$$

And let $g(a, b) = (g^J(a, b))_{J \in \mathcal{F}}$ be a vector-valued function, then $g(a, b)$ satisfies the following Pfaffian equations [3, Theorem 22] :

$$\partial_{a_{ij}} g^J = \sum_{k=1}^n a_{ik} \partial_{b_k} \partial_{b_j} g^J \quad (1 \leq i \leq d, 1 \leq j \leq n, J \in \mathcal{F}), \quad (5)$$

$$\partial_{b_j} g^J = g^{J \cup \{j\}} \quad (j \in J^c, J \in \mathcal{F}), \quad (6)$$

$$\partial_{b_j} g^J = - \sum_{k \in J} \alpha_J^{jk}(a) \left(b_k g^J + \sum_{\ell \in J^c} \alpha_{k\ell}(a) g^{J \cup \ell} \right) \quad (j \in J, J \in \mathcal{F}). \quad (7)$$

Here, $(\alpha_F^{ij}(a))_{i,j \in F}$ is the inverse matrix of $\alpha_F(a) = \left(\sum_{k=1}^d a_{ki} a_{kj} \right)_{i,j \in F}$, which is a submatrix of the gram matrix of a . Note that the right hand side of (5) can be rewritten, with recourse to (6) and (7), as a linear combination of g^J with rational functions as coefficients.

3 Inclusion-Exclusion Identity for Faces

Let P be the polyhedron defined by (1), and suppose the family of the bounding half-spaces for P is given by (2). Then, we have the following inclusion–exclusion identity [1]:

$$\prod_{j=1}^n H \left(\sum_{i=1}^d \tilde{a}_{ij} x_i + \tilde{b}_j \right) = \sum_{J \in \mathcal{F}} \prod_{j \in J} \left(H \left(\sum_{i=1}^d \tilde{a}_{ij} x_i + \tilde{b}_j \right) - 1 \right) \quad (x \in \mathbf{R}^d). \quad (8)$$

In [3], we showed that if the polyhedron P is in general position, there exists a neighborhood U of the parameter $(\tilde{a}, \tilde{b}) \in \mathbf{R}^{d \times n} \times \mathbf{R}^n$ such that

$$\prod_{j=1}^n H \left(\sum_{i=1}^d a_{ij} x_i + b_j \right) = \sum_{J \in \mathcal{F}} \prod_{j \in J} \left(H \left(\sum_{i=1}^d a_{ij} x_i + b_j \right) - 1 \right) \quad (9)$$

holds for any $(a, b) \in U$ and $x \in \mathbf{R}^d$. The left hand sides of (8) and (9) are the indicator functions for the corresponding polyhedra. In this section we give analogous identities for the indicator functions of a face of the polyhedra.

Let \mathcal{F} be the abstract simplicial complex for the polyhedron P . For $J \in \mathcal{F}$,

$$\mathcal{F}_J := \{F \in \mathcal{F} \mid J \subset F\}.$$

For parameter $(a, b) \in \mathbf{R}^{d \times n} \times \mathbf{R}^n$ and $J \in \mathcal{F}$, we define a hyperplane $V(J, a, b)$ by

$$V(J, a, b) = \left\{ x \in \mathbf{R}^d \mid \sum_{i=1}^d a_{ij} x_i + b_j = 0 \ (j \in J) \right\}. \quad (10)$$

Proposition 1. *Suppose the polyhedron P is in general position. For each $J \in \mathcal{F}$, we have the equation*

$$\prod_{j \in [n] \setminus J} H \left(\sum_{i=1}^d \tilde{a}_{ij} x_i + \tilde{b}_j \right) = \sum_{F \in \mathcal{F}_J} \prod_{j \in F \setminus J} \left(H \left(\sum_{i=1}^d \tilde{a}_{ij} x_i + \tilde{b}_j \right) - 1 \right).$$

for any $x \in V(J, \tilde{a}, \tilde{b})$.

Proof. Let s be the number of elements in the set J . Since the polyhedron P is in general position, we have $s \leq d$. By replacing the indices, we can assume $J = \{n - s + 1, \dots, n\}$ without loss of generality. Applying the Euclidean transformation for P , we can assume $\tilde{a}_{ij} = 0, \tilde{b}_j = 0$ ($1 \leq i \leq d - s, n - s + 1 \leq$

$j \leq n$). Then, by the assumption of general position, the vectors $\tilde{a}_{n-s+1}, \dots, \tilde{a}_n$ are linearly independent (see, [3, Corollary 20]). Hence we have

$$V(J, \tilde{a}, \tilde{b}) = \{x \in \mathbf{R}^d \mid x_{d-s+1} = \dots = x_d = 0\},$$

and the problem is reduced to the proof of

$$\prod_{j=1}^{n-s} H\left(\sum_{i=1}^{d-s} \tilde{a}_{ij} y_i + \tilde{b}_j\right) = \sum_{F \in \mathcal{F}_J} \prod_{j \in F \setminus J} \left(H\left(\sum_{i=1}^{d-s} \tilde{a}_{ij} y_i + \tilde{b}_j\right) - 1\right) \quad (11)$$

for arbitrary $y = (y_1, \dots, y_{d-s})^\top \in \mathbf{R}^{d-s}$. Let us consider a polyhedron

$$P' := \left\{ y \in \mathbf{R}^{d-s} \mid \sum_{i=1}^{d-s} \tilde{a}_{ij} y_i + \tilde{b}_j \geq 0, 1 \leq j \leq n-s \right\}.$$

Suppose that there are t redundant inequalities in the definition of P' . By replacing indices, we can assume that the redundant inequalities are $\sum_{i=1}^{d-s} \tilde{a}_{ij} y_i + \tilde{b}_j \geq 0$, $(n-s-t+1 \leq j \leq n-s)$. Then, all facets of P' are given by

$$F'_j := P' \cap \left\{ y \in \mathbf{R}^{d-s} \mid \sum_{i=1}^{d-s} \tilde{a}_{ij} y_i + \tilde{b}_j = 0 \right\} \quad (1 \leq j \leq n-s-t),$$

and the abstract simplicial complex for P' is

$$\mathcal{F}' = \left\{ J' \subset \{1, 2, \dots, n-s-t\} \mid \bigcap_{j \in J'} F'_j \neq \emptyset \right\}.$$

Applying the inclusion-exclusion identity for P' , we have

$$\prod_{j=1}^{n-s-t} H\left(\sum_{i=1}^{d-s} \tilde{a}_{ij} y_i + \tilde{b}_j\right) = \sum_{F \in \mathcal{F}'} \prod_{j \in F} \left(H\left(\sum_{i=1}^{d-s} \tilde{a}_{ij} y_i + \tilde{b}_j\right) - 1\right). \quad (12)$$

Since the left hand sides of (11) and (12) are both equal to the indicator function of P' , they are equal to each other. Consequently, we need to show that the right hand sides of (11) and (12) are equal. It is easy to show that the mapping

$$\psi: \mathcal{F}' \rightarrow \mathcal{F}_J \quad (J' \mapsto J \cup J')$$

is a bijection. Rewriting the right hand side of (12) in terms of \mathcal{F}_J , we have the same expression for the right hand side of (11). \square

For use in the next section, we extend Proposition 1. We introduce the

following notation. For parameters $(a, b) \in \mathbf{R}^{d \times n} \times \mathbf{R}^n$, let

$$\begin{aligned} H_j(a, b) &= \left\{ x \in \mathbf{R}^d \mid \sum_{i=1}^d a_{ij}x_i + b_j \geq 0 \right\} \quad (1 \leq j \leq n), \\ \mathcal{H}(a, b) &= \{\mathcal{H}_1(a, b), \dots, \mathcal{H}_n(a, b)\}, \\ P(a, b) &= \bigcap_{j=1}^n H_j(a, b), \\ F_j(a, b) &= \left\{ x \in \mathbf{R}^d \mid \sum_{i=1}^d a_{ij}x_i + b_j = 0 \right\} \cap P(a, b), \\ \mathcal{F}(a, b) &= \left\{ J \subset [n] \mid \bigcap_{j \in J} F_j(a, b) \neq \emptyset \right\}. \end{aligned}$$

The following lemma holds.

Lemma 1. *Suppose the polyhedron P is in general position. Then, there exists a neighborhood U of the parameter $(\tilde{a}, \tilde{b}) \in \mathbf{R}^{d \times n} \times \mathbf{R}^n$ which satisfies the following: for any parameter (a, b) in U , $P(a, b)$ is in general position and $\mathcal{F}(a, b) = \mathcal{F}$ holds.*

Proof. We put $a_{i0} = 0$ ($1 \leq i \leq d$), $b_0 = 0$, and

$$\begin{aligned} \hat{F}_j(a, b) &= \left\{ (x_0, x) \in \mathbf{R} \times \mathbf{R}^d \mid \begin{array}{l} \sum_{i=1}^d a_{ij}x_i + b_j = 0, \\ \sum_{i=1}^d a_{ik}x_i + b_k \geq 0 \\ (0 \leq k \leq n) \end{array} \right\} \quad (0 \leq j \leq n) \\ \hat{\mathcal{F}}(a, b) &= \left\{ J \subset \{0, 1, \dots, n\} \mid \bigcap_{j \in J} \hat{F}_j(a, b) \neq \emptyset \right\}. \end{aligned}$$

By Theorem 23 in [3], the set

$$U := \left\{ (a, b) \in \mathbf{R}^{d \times n} \times \mathbf{R}^n \mid \begin{array}{l} P(a, b) \text{ is in general position,} \\ \hat{\mathcal{F}}(a, b) = \hat{\mathcal{F}}(\tilde{a}, \tilde{b}) \end{array} \right\}$$

is a neighborhood of the point $(\tilde{a}, \tilde{b}) \in \mathbf{R}^{d \times n} \times \mathbf{R}^n$. Consider arbitrary $(a, b) \in U$, from Corollary 19 in [3], we have $\mathcal{F}(a, b) = \mathcal{F}(\tilde{a}, \tilde{b}) = \mathcal{F}$ from $\hat{\mathcal{F}}(a, b) = \hat{\mathcal{F}}(\tilde{a}, \tilde{b})$. By Lemma 22 in [3], all facets of $P(a, b)$ are given by $F_j(a, b)$ ($\{j\} \in \mathcal{F}(a, b)$). The equation $\mathcal{F}(a, b) = \mathcal{F}$ implies $\{j\} \in \mathcal{F}(a, b)$ for all $j \in [n]$. Consequently, $\mathcal{H}(a, b)$ is the bounding half-spaces for $P(a, b)$ and $P(a, b)$ is in general position. \square

Finally, we have the following.

Theorem 1. Suppose the polyhedron P is in general position and $J \in \mathcal{F}$. There exists a neighborhood U of $(\tilde{a}, \tilde{b}) \in \mathbf{R}^{d \times n} \times \mathbf{R}^n$ such that for any $(a, b) \in U$ and $x \in V(J, a, b)$, we have

$$\prod_{j \in [n] \setminus J} H \left(\sum_{i=1}^d a_{ij} x_i + b_j \right) = \sum_{F \in \mathcal{F}_J} \prod_{j \in F \setminus J} \left(H \left(\sum_{i=1}^d a_{ij} x_i + b_j \right) - 1 \right). \quad (13)$$

Proof. Let U be a neighborhood of (\tilde{a}, \tilde{b}) in Lemma 1. Then the polyhedron $P(a, b)$ is in general position. By Lemma 22 in [3], the abstract simplicial complex associated with $P(a, b)$ is equivalent to $\mathcal{F}(a, b)$. The equation $\mathcal{F}(a, b) = \mathcal{F}$ implies $J \in \mathcal{F}(a, b)$. Hence, we can apply Proposition 1 which gives

$$\prod_{j \in [n] \setminus J} H \left(\sum_{i=1}^d a_{ij} x_i + b_j \right) = \sum_{F \in \mathcal{F}_J(a, b)} \prod_{j \in F \setminus J} \left(H \left(\sum_{i=1}^d a_{ij} x_i + b_j \right) - 1 \right)$$

for any $x \in V(J, a, b)$. Here, we put $\mathcal{F}_J(a, b) = \{F \subset \mathcal{F}(a, b) \mid J \subset F\}$. Since $\mathcal{F}(a, b) = \mathcal{F}$ implies $\mathcal{F}_J(a, b) = \mathcal{F}_J$, we thus have equation (13). \square

4 Derivatives of the Probability Content

In this section we derive an expression for the function $g^J(a, b)$ ($J \in \mathcal{F}$), which is a derivative of the function $\varphi(a, b)$ defined by the probability content of the polyhedron P for the multivariate normal distribution. We then show that the function $g^J(a, b)$ can be expressed as an integral on the hyperplane (10). For simplicity, we put

$$\varphi_F(a, b) = \int_{\mathbf{R}^d} \exp \left(-\frac{1}{2} \sum_{i=1}^d x_i^2 \right) \prod_{j \in F} H(-f_j(a, b, x)) dx. \quad (14)$$

Then, the function $\varphi(a, b)$ in (3) can be written as

$$\varphi(a, b) = \sum_{F \in \mathcal{F}} \frac{(-1)^{|F|}}{(2\pi)^{d/2}} \varphi_F(a, b).$$

In order to obtain expressions for the function $g^J(a, b)$, we first consider $\partial_b^J \bullet \varphi_F(a, b)$ for $F \in \mathcal{F}$. Consider the following case:

Lemma 2. Let $1 \leq p \leq q \leq d$, $J = \{1, \dots, p\}$, and $F = \{1, \dots, p, \dots, q\}$. Suppose that parameter a satisfies $a_{ij} = 0$ ($p < i \leq d$, $1 \leq j \leq p$) and $\alpha_F(a) =$

$\left(\sum_{k=1}^d a_{ki}a_{kj}\right)_{i,j \in F}$ is a non-singular matrix, i.e.,

$$a = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1p} & a_{1(p+1)} & \cdots & a_{1q} & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} & a_{p(p+1)} & \cdots & a_{pq} & * & \cdots & * \\ 0 & \cdots & 0 & a_{(p+1)(p+1)} & \cdots & a_{(p+1)q} & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots & & & \vdots \\ 0 & \cdots & 0 & a_{d(p+1)} & \cdots & a_{dq} & * & \cdots & * \end{pmatrix}$$

and the vectors $a_1, \dots, a_p, \dots, a_q$ are linearly independent. Then, the function $\partial_b^J \bullet \varphi_F(a, b)$ is equal to the integral

$$\frac{(-1)^{|J|}}{\sqrt{|\alpha_J(a)|}} \int_{V(J, a, b)} \exp\left(-\frac{1}{2} \sum_{i=1}^d x_i^2\right) \prod_{j \in F \setminus J} H(-f_j(a, b, x)) \mu(dx). \quad (15)$$

Here, μ is the volume element of the hyperplane $V(J, a, b)$.

Proof. Let a $d \times d$ matrix $U = (u_{ij})$ be

$$U = \begin{pmatrix} a_{11} & \cdots & a_{1p} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & & \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & 0 & & & 1 \end{pmatrix}.$$

The elements of U can be written as

$$u_{ij} = \begin{cases} a_{ij} & (1 \leq j \leq p) \\ \delta_{ij} & (p < j \leq d) \end{cases}.$$

Here, δ_{ij} is Kronecker's delta. As the vectors a_1, \dots, a_p are linearly independent, the matrix U is regular, and we have $|U|^2 = |\alpha_J(a)|$. We denote the inverse matrix of U by $U^{-1} = (u^{ij})$. Consider a transformation of variables $y_j = \sum_{i=1}^d u_{ij}x_i$ ($1 \leq j \leq d$). By the relationships

$$x_i = \begin{cases} \sum_{k=1}^p u^{ki}y_k & (1 \leq i \leq p) \\ y_i & (p+1 \leq i \leq d) \end{cases},$$

$$y_j = \sum_{i=1}^d a_{ij}x_i + b_j \quad (1 \leq j \leq p),$$

the integral $\varphi_F(a, b)$ can be written as

$$\begin{aligned} & \frac{1}{\sqrt{|\alpha_J(a)|}} \int_{\mathbf{R}^d} e^{-\frac{1}{2} \sum_{i=1}^p (\sum_{k=1}^p u^{ki} y_k)^2 - \frac{1}{2} \sum_{i=p+1}^d y_i^2} \prod_{j=1}^p H(-y_j - b_j) \\ & \times \prod_{j=p+1}^q H\left(-\sum_{i=1}^p \sum_{k=1}^p a_{ij} u^{ki} y_k - \sum_{i=p+1}^d a_{ij} y_i - b_j\right) dy_1 \dots dy_d \\ & = \int_{-\infty}^{-b_1} \dots \int_{-\infty}^{-b_p} G(y_1, \dots, y_p; a, b, U) dy_p \dots dy_1. \end{aligned}$$

Here, we put

$$\begin{aligned} & G(y_1, \dots, y_p; a, b, U) \\ & = \frac{e^{-\frac{1}{2} \sum_{i=1}^p (\sum_{k=1}^p u^{ki} y_k)^2}}{\sqrt{|\alpha_J(a)|}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \sum_{i=p+1}^d y_i^2} \\ & \times \prod_{j=p+1}^q H\left(-\sum_{i=1}^p \sum_{k=1}^p a_{ij} u^{ki} y_k - \sum_{i=p+1}^d a_{ij} y_i - b_j\right) dy_{p+1} \dots dy_d. \end{aligned}$$

Since $G(y_1, \dots, y_p; a, b, U)$ is a continuous function with respect to y_1, \dots, y_p , we have $\partial_b^J \bullet \varphi_F(a, b) = (-1)^p G(-b_1, \dots, -b_p; a, b, U)$.

When $p = d$, we have

$$\partial_b^J \bullet \varphi_F(a, b) = \frac{(-1)^d}{\sqrt{|\alpha_J(a)|}} e^{-\frac{1}{2} \sum_{i=1}^d (\sum_{k=1}^d u^{ki} b_k)^2}$$

Since $-(U^{-1})^\top b$ is the unique point in $V(J, a, b)$, $\partial_b^J \bullet \varphi_F(a, b)$ equals to (15).

Suppose $p \neq d$, and define a mapping $\psi(x) = (y_{p+1}, \dots, y_d)$ for the hyper-plane $V(J, a, b)$ to \mathbf{R}^{d-p} by $y_j = x_j$ ($p+1 \leq j \leq d$), then this mapping is a local coordinate system on $V(J, a, b)$. At this coordinate, the functions x_1, \dots, x_d on $V(J, a, b)$ can be written as

$$x_i = \begin{cases} -\sum_{k=1}^p u^{ki} b_k & (1 \leq i \leq p), \\ y_i & (p+1 \leq i \leq d). \end{cases}$$

The Riemannian metric induced on $V(J, a, b)$ is $\sum_{i=1}^d dx_i \otimes dx_i = \sum_{j=p+1}^d dy_j \otimes dy_j$. Calculating (15) with this coordinate, we have $g^J(a, b) = (-1)^p G(-b_1, \dots, -b_p; a, b, U)$. \square

We now extend this lemma.

Lemma 3. *Let $1 \leq p \leq q \leq d$, $J = \{1, \dots, p\}$, and $F = \{1, \dots, p, \dots, q\}$. Suppose $\alpha_F(a)$ is a regular matrix. Then, the function $\partial_b^J \bullet \varphi_F(a, b)$ is equal to (15).*

Proof. It is sufficient to show that this reduces to Lemma 2.

For a suitable special orthogonal matrix R , the $d \times n$ matrix $a' := Ra$ satisfies the condition $a'_{ij} = 0$ ($p < i \leq d$, $1 \leq j \leq p$). Since $\alpha_F(a) = \alpha_F(a')$, $\alpha_F(a')$ is also a regular matrix by this assumption. Hence, the parameter (a', b) satisfies the condition of Lemma 2.

Since the Lebesgue measure is invariant under the action of the special orthogonal group, we have $\varphi_F(a, b) = \varphi_F(a', b)$ for any $b \in \mathbf{R}^n$. Consequently, we have $\partial_b^J \bullet \varphi_F(a, b) = \partial_b^J \bullet \varphi_F(a', b)$, and $1/\sqrt{|\alpha_J(a)|} = 1/\sqrt{|\alpha_J(a')|}$.

Considering (15), we put

$$\tilde{\varphi}_F(a, b) = \int_{V(J, a, b)} \exp\left(-\frac{1}{2} \sum_{i=1}^d x_i^2\right) \prod_{j \in F \setminus J} H(-f_j(a, b, x)) \mu(dx).$$

We need to show $\tilde{\varphi}_F(a, b) = \tilde{\varphi}_F(a', b)$. When $p = d$, this relation is trivial. Suppose that $p < d$. Take vectors $v_j = (v_{1j}, \dots, v_{dj})^\top$ ($q+1 \leq j \leq d$) such that $a_1, \dots, a_q, v_{q+1}, \dots, v_d$ are linearly independent. Let $U = (u_{ij})$ be a matrix obtained by arranging these vectors,

$$u_{ij} = \begin{cases} a_{ij} & (1 \leq j \leq q), \\ v_{ij} & (q+1 \leq j \leq d). \end{cases}$$

We denote the inverse matrix of U , by $U^{-1} = (u^{ij})$. And define a matrix $U' = (u'_{ij})$ as $U' = RU$, and denote its inverse by $U'^{-1} = (u'^{ij})$.

First, we calculate $\tilde{\varphi}_F(a, b)$. If we define a map $\psi(x) = (y_{p+1}, \dots, y_d)$ from the hyperplane $V(J, a, b)$ to \mathbf{R}^{d-p} by $y_j = \sum_{i=1}^d u_{ij} x_i$ ($p+1 \leq j \leq d$), then it is a local coordinate system on $V(J, a, b)$. With this coordinate, the function x_i on $V(J, a, b)$ can be written as

$$x_i = -\sum_{k=1}^p u^{ki} b_k + \sum_{k=p+1}^d u^{ki} y_k \quad (1 \leq i \leq d).$$

Hence, the Riemannian metric on the hyperplane $V(J, a, b)$ is

$$\sum_{i=1}^d dx_i \otimes dx_i = \sum_{k=1}^p \sum_{\ell=1}^p \left(\sum_{i=1}^d u^{ki} u^{\ell i} \right) dy_k \otimes dy_\ell.$$

Let D be the determinant of the matrix $\left(\sum_{i=1}^d u^{ki} u^{\ell i} \right)_{1 \leq k, \ell \leq p}$. The integral $\tilde{\varphi}_F(a, b)$ can be written as

$$\frac{1}{\sqrt{|D|}} \int_{\mathbf{R}^{d-p}} e^{-\frac{1}{2} \sum_{i=1}^d (-\sum_{k=1}^p u^{ki} b_k + \sum_{k=p+1}^d u^{ki} y_k)^2} \prod_{j=p+1}^q H(-y_j - b_j) \prod_{j=p+1}^d dy_j.$$

Next, we calculate $\tilde{\varphi}_F(a', b)$. If we define a map $\psi'(x) = (y_{p+1}, \dots, y_d)$ from the hyperplane $V(J, a', b)$ to \mathbf{R}^{d-p} by $y_j = \sum_{i=1}^d u'_{ij} x_i$ ($p+1 \leq j \leq d$), then it

is a local coordinate system on $V(J, a', b)$. With this coordinate, the function x_i on $V(J, a', b)$ can be written as

$$x_i = - \sum_{k=1}^p u'^{ki} b_k + \sum_{k=p+1}^d u'^{ki} y_k \quad (1 \leq i \leq d).$$

By $U'^{-1} = U^{-1} R^\top$, the Riemannian metric on $V(J, a', b)$ is

$$\sum_{i=1}^d dx_i \otimes dx_i = \sum_{k=1}^p \sum_{\ell=1}^p \left(\sum_{i=1}^d u'^{ki} u'^{\ell i} \right) dy_k \otimes dy_\ell = \sum_{k=1}^p \sum_{\ell=1}^p \left(\sum_{i=1}^d u^{ki} u^{\ell i} \right) dy_k \otimes dy_\ell.$$

And we have

$$\begin{aligned} \sum_{i=1}^d x_i^2 &= \sum_{i=1}^d \left(- \sum_{k=1}^p u'^{ki} b_k + \sum_{k=p+1}^d u'^{ki} y_k \right)^2 \\ &= \sum_{i=1}^d \left(- \sum_{k=1}^p u^{ki} b_k + \sum_{k=p+1}^d u^{ki} y_k \right)^2. \end{aligned}$$

Hence we have $\tilde{\varphi}_F(a', b) = \tilde{\varphi}_F(a, b)$. \square

From Lemma 3, we have the following.

Lemma 4. *Let $F \in \mathcal{F}$ and suppose $\alpha_F(a)$ is a regular matrix. Then, we have*

$$\begin{aligned} &\partial_b^J \bullet \varphi_F(a, b) \\ &= \begin{cases} \frac{(-1)^{|J|}}{\sqrt{|\alpha_J(a)|}} \int_{V(J, a, b)} e^{-\frac{1}{2} \sum_{i=1}^d x_i^2} \prod_{j \in F \setminus J} H(-f_j(a, b, x)) \mu(dx) & (J \subset F), \\ 0 & (J \not\subset F). \end{cases} \end{aligned}$$

Proof. When $J \subset F$, it reduces to Lemma 3 since we can assume $J = \{1, \dots, p\}$, $F = \{1, \dots, p, \dots, q\}$, and $1 \leq p \leq q \leq d$ without loss of generality. When $J \not\subset F$, the integral in (14) does not depend on the variables b_j ($j \in J \setminus F$). Hence, the derivative with respect to b_j is 0. \square

Theorem 2. *Suppose that the polyhedron P is in general position and $J \in \mathcal{F}$, then there exists a neighborhood U of the parameter $(\tilde{a}, \tilde{b}) \in \mathbf{R}^{d \times n} \times \mathbf{R}^n$ such that the equation*

$$g^J(a, b) = \frac{1}{(2\pi)^{d/2} \sqrt{|\alpha_J(a)|}} \int_{V(J, a, b)} e^{-\frac{1}{2} \sum_{i=1}^d x_i^2} \prod_{j \in [n] \setminus J} H(f_j(a, b, x)) \mu(dx) \quad (16)$$

holds for any $(a, b) \in U$.

Proof. By (3) and (4), we have

$$(2\pi)^{d/2} g^J(a, b) = \sum_{F \in \mathcal{F}} \partial_b^J \bullet (-1)^{|F|} \int_{\mathbf{R}^d} e^{-\frac{1}{2} \sum_{i=1}^d x_i^2} \prod_{j \in F} H(-f_j(a, b, x)) dx.$$

Applying Lemma 4 to each term on the right hand side of the above equation, we can show that $(2\pi)^{d/2} g^J(a, b)$ is equal to

$$\begin{aligned} & \frac{1}{\sqrt{|\alpha_J(a)|}} \int_{V(J, a, b)} e^{-\frac{1}{2} \sum_{i=1}^d x_i^2} \sum_{F \in \mathcal{F}_J} (-1)^{|F \setminus J|} \prod_{j \in F \setminus J} H(-f_j(a, b, x)) \mu(dx) \\ &= \frac{1}{\sqrt{|\alpha_J(a)|}} \int_{V(J, a, b)} e^{-\frac{1}{2} \sum_{i=1}^d x_i^2} \sum_{F \in \mathcal{F}_J} \prod_{j \in F \setminus J} H(f_j(a, b, x) - 1) \mu(dx). \end{aligned}$$

From Theorem 1, we have the equation (16). \square

5 Holonomic Gradient Method

In this section, we discuss the computation of the probability content of a polyhedron with a multivariate normal distribution for the case where the polyhedron is in general position and bounded, and the case where the polyhedron is a simplicial cone.

5.1 The Bounded Case

Let us consider the case where the polyhedron P in general position is bounded.

Lemma 5. *Suppose the polyhedron P is bounded. Then, the set*

$$\left\{ x \in \mathbf{R}^d \mid \sum_{i=1}^d \tilde{a}_{ij} x_i \geq 0 \ (1 \leq j \leq n) \right\} \quad (17)$$

contain only the origin.

Proof. By Proposition 1.12 in [12], the set (17) is equal to

$$\{ y \in \mathbf{R}^d \mid x + ty \in P \ (x \in P, t \geq 0) \}.$$

Since P is bounded, this set does not contain any element except the origin. \square

Proposition 2. *Suppose the polyhedron P in general position is bounded. Then, for $J \in \mathcal{F}$, we have*

$$g^J(\tilde{a}, 0) = \begin{cases} \frac{1}{\sqrt{|\alpha_J(\tilde{a})|}} & (|J| = d) \\ 0 & (else) \end{cases}. \quad (18)$$

Proof. Calculating the left hand side, we have

$$\begin{aligned}
& (2\pi)^{d/2} g^J(\tilde{a}, 0) \\
&= \lim_{t \rightarrow +0} (2\pi)^{d/2} g^J(\tilde{a}, tb) \\
&= \lim_{t \rightarrow +0} \frac{1}{\sqrt{|\alpha_J(\tilde{a})|}} \int_{V(J, \tilde{a}, tb)} e^{-\frac{1}{2} \sum_{i=1}^d x_i^2} \prod_{j \in [n] \setminus J} H(f_j(\tilde{a}, tb, x)) \mu(dx) \\
&= \frac{1}{\sqrt{|\alpha_J(\tilde{a})|}} \int_{V(J, \tilde{a}, 0)} e^{-\frac{1}{2} \sum_{i=1}^d x_i^2} \prod_{j \in [n] \setminus J} H(f_j(\tilde{a}, 0, x)) \mu(dx).
\end{aligned}$$

By Lemma 5, the integral domain is $\{0\}$. Hence we have (18). \square

Consequently, in order to compute the probability content of P for a multivariate normal distribution, we can take the path of the HGM as

$$a(t) = \tilde{a}, b(t) = \tilde{t}\tilde{b} \quad (0 \leq t \leq 1).$$

This path does not pass through the singular locus of the Pfaffian equations (5), (6) and (7). The initial value $g^J(a(0), 0)$ at $t = 0$ is given explicitly by (18).

5.2 The Simplicial Cone Case

Consider the case where the polyhedron P is a simplicial cone, i.e., $n = d$ and the vectors $\tilde{a}_1, \dots, \tilde{a}_d$ are linearly independent. We can assume without loss of generality that \tilde{a} is an upper triangular matrix. Then define $\gamma(t) = (a(t), b(t))$ by

$$a(t) = (1 - t)\text{diag}(\tilde{a}_{11}, \dots, \tilde{a}_{dd}) + t\tilde{a}, b(t) = \tilde{t}\tilde{b} \quad (0 \leq t \leq 1).$$

This does not pass through the singular locus of the Pfaffian equation. The initial value is

$$g^J(a(0), b(0)) = \frac{1}{\left| \prod_{j \in J} \tilde{a}_{jj} \right|} \sqrt{\frac{\pi}{2}}^{d-|J|}.$$

6 Numerical Experiments

In this section we compare the performance of our HGM method with a Monte Carlo simulation method. In the Monte Carlo simulation method, we used the computer system R [10].

First, we evaluate the probability contents of simplexes. For an integer $d \geq 2$, we define polyhedra P_d and Q_d as

$$\begin{aligned}
P_d &= \left\{ x \in \mathbf{R}^d \mid \begin{array}{l} x_i + \frac{\sqrt{d}}{2} \geq 0 \ (1 \leq i \leq d), \\ -x_1 - \dots - x_d + \frac{\sqrt{d}}{2} \geq 0 \end{array} \right\}, \\
Q_d &= \left\{ x \in \mathbf{R}^d \mid \begin{array}{l} x_i - \frac{\sqrt{d}}{2} \geq 0 \ (1 \leq i \leq d), \\ -x_1 - \dots - x_d + \frac{(2d+1)\sqrt{d}}{2} \geq 0 \end{array} \right\}
\end{aligned}$$

Both P_d and Q_d are simplexes, and they are in general position and bounded. In the Monte Carlo method, we generated 1,000,000 simulations from a normal distribution and computed the fraction of samples that fell into simplexes.

The probability contents obtained by the HGM and Monte Carlo methods are given in Tables 1 and 2. We also show the computational times for the HGM in the tables.

d	HGM	time of HGM(s)	MC
2	0.285205	0.00	0.2849
3	0.251995	0.00	0.2493
4	0.241744	0.01	0.2429
5	0.242724	0.02	0.2428
6	0.250219	0.09	0.2394
7	0.261920	0.32	0.2572
8	0.276510	1.04	0.2787
9	0.293138	3.15	0.2859
10	0.311198	9.51	0.3072

Table 1: The probability content of P_d as obtained by the HGM and Monte Carlo methods.

d	HGM	time of HGM(s)	MC
2	5.1758e-02	0.00	5.1917e-02
3	7.0235e-03	0.00	7.0850e-03
4	6.3101e-04	0.00	6.0400e-04
5	3.9722e-05	0.02	5.5000e-05
6	1.8042e-06	0.10	3.0000e-06
7	5.9878e-08	0.30	0.0000e+00
8	1.4799e-09	0.85	0.0000e+00
9	1.1393e-11	2.25	0.0000e+00
10	1.2861e-11	5.74	0.0000e+00

Table 2: The probability content of Q_d as obtained by the HGM and Monte Carlo methods.

Note that the accuracy of the Monte Carlo method is low when the probability content of Q_d is very small, and for dimensions greater than 6, the Monte Carlo method could not evaluate the probability content of Q_d . The number of samples was not enough to evaluate the probability.

Next, we estimate the probability content of a simplicial cone. For an integer

$d \geq 2$, we define a polyhedron C_d as

$$C_d = \left\{ x \in \mathbf{R}^d \mid \sum_{i=1}^d a_{ij}x_i + \frac{\sqrt{d}}{2} \geq 0 \ (1 \leq j \leq d) \right\},$$

$$a_{ij} = \begin{cases} (i+j)/100 & (i < j) \\ 1 & (i = j) \\ 0 & (i > j) \end{cases}.$$

We can evaluate the multivariate normal probability of a simplicial cone using the method presented in Subsection 5.2. Table 3 shows the probability content of C_d evaluated by the HGM and Monte Carlo methods, and the computational times for the HGM. In the Monte Carlo method, we generated 1,000,000 samples.

dim	HGM	time of HGM(s)	MC
2	0.580822	0.00	0.5813
3	0.532131	0.01	0.5331
4	0.512854	0.05	0.5189
5	0.509868	0.52	0.5216
6	0.516602	3.98	0.5184
7	0.529243	25.90	0.5374
8	0.545340	147.00	0.5488
9	0.563203	770.00	0.5606
10	0.581630	3811.00	0.5691

Table 3: Results of HGM for multivariate normal probabilities of simplicial cones

References

- [1] H. Edelsbrunner. The union of balls and its dual shape. *Discrete & Computational Geometry*, 13:415–440, 1995.
- [2] H. Hashiguchi, Y. Numata, N. Takayama, and A. Takemura. Holonomic gradient method for the distribution function of the largest root of Wishart matrix. *Journal of Multivariate Analysis*, 117:296–312, 2013.
- [3] T. Koyama. Holonomic modules associated with multivariate normal probabilities of polyhedra. <http://arxiv.org/abs/1311.6905>, 2013.
- [4] T. Koyama and A. Takemura. Calculation of orthant probabilities by the holonomic gradient method. <http://arxiv.org/abs/1211.6822>, 2012.

- [5] T. Koyama and A. Takemura. Holonomic gradient method for distribution function of a weighted sum of noncentral chi-square random variables. <http://arxiv.org/abs/1503.00378>, 2015.
- [6] Tamio Koyama, Hiromasa Nakayama, Katsuyoshi Ohara, Tomonari Sei, and Nobuki Takayama. Software packages for holonomic gradient method. In Hoon Hong and Chee Yap, editors, *Mathematical Software – ICMS 2014*, volume 8592 of *Lecture Notes in Computer Science*, pages 706–712. Springer Berlin Heidelberg, 2014.
- [7] D. Q. Naiman and H. P. Wynn. Inclusion–exclusion–bonferroni identities and inequalities for discrete tube–like problems via euler characteristics. *The Annals of Statistics*, 20(1):43–76, 1992.
- [8] H. Nakayama, K. Nishiyama, M. Noro, K. Ohara, T. Sei, N. Takayama, and A. Takemura. Holonomic gradient descent and its application to the Fisher-Bingham integral. *Advances in Applied Mathematics*, 47:639–658, 2011.
- [9] T. Oaku. Computation of the characteristic variety and the singular locus of a system of differential equations with polynomial coefficients. *Japan Journal of Industrial and Applied Mathematics*, 11:485–497, 1994.
- [10] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0.
- [11] Grady Weyenberg, Daniel K Howe, and Ruriko Yoshisa. Normalizing kernels in the billera-holmes-vogtmann treespace. <http://arxiv.org/pdf/1506.00142v1>, 2015.
- [12] G. M. Ziegler. *Lectures on Polytopes*. Springer-Verlag, New York, 1995.